Stochastic Link And Group Detection

The goal of this research is to apply a probabilistic approach to discover and identify underlying structure (e.g. groups). For example we may want to extract groups of people who are working together from a log of all meetings that occurred in a city during a given time. With such groups, we could ask questions such as: list all members in a group, list all groups for which entity 1 and entity 2 are members, determine the missing entity given a partial link, determine if two entities are the same (aliases) or if one entity is two objects. The model takes the form of a Bayes network that captures the demographic data about entities and link information about entity transactions, meeting, etc. The model then applies probabilistic scores generated by a Bayes classifier and uses a heuristic optimization technique to predict group relationships.

Big Noisy Data-base of Links

Record # 456621:

Doe and Smith were booked in the same hotel on 3/6/93

Before..

There must be something useful here..but I'm drowning in noise

Analyst

Potential coll-aborators of Doe:

The pattern of travel among the group consisting of (Doe, Smith, Jones, Brown) during 93-96 can't be explained by coincidence

..After

Big Noisy Data-base of Links

Our Algorithms

This is more useful objective information